OIPE
AUG 04 2005
PATENT & TRADEMARK OFFICE

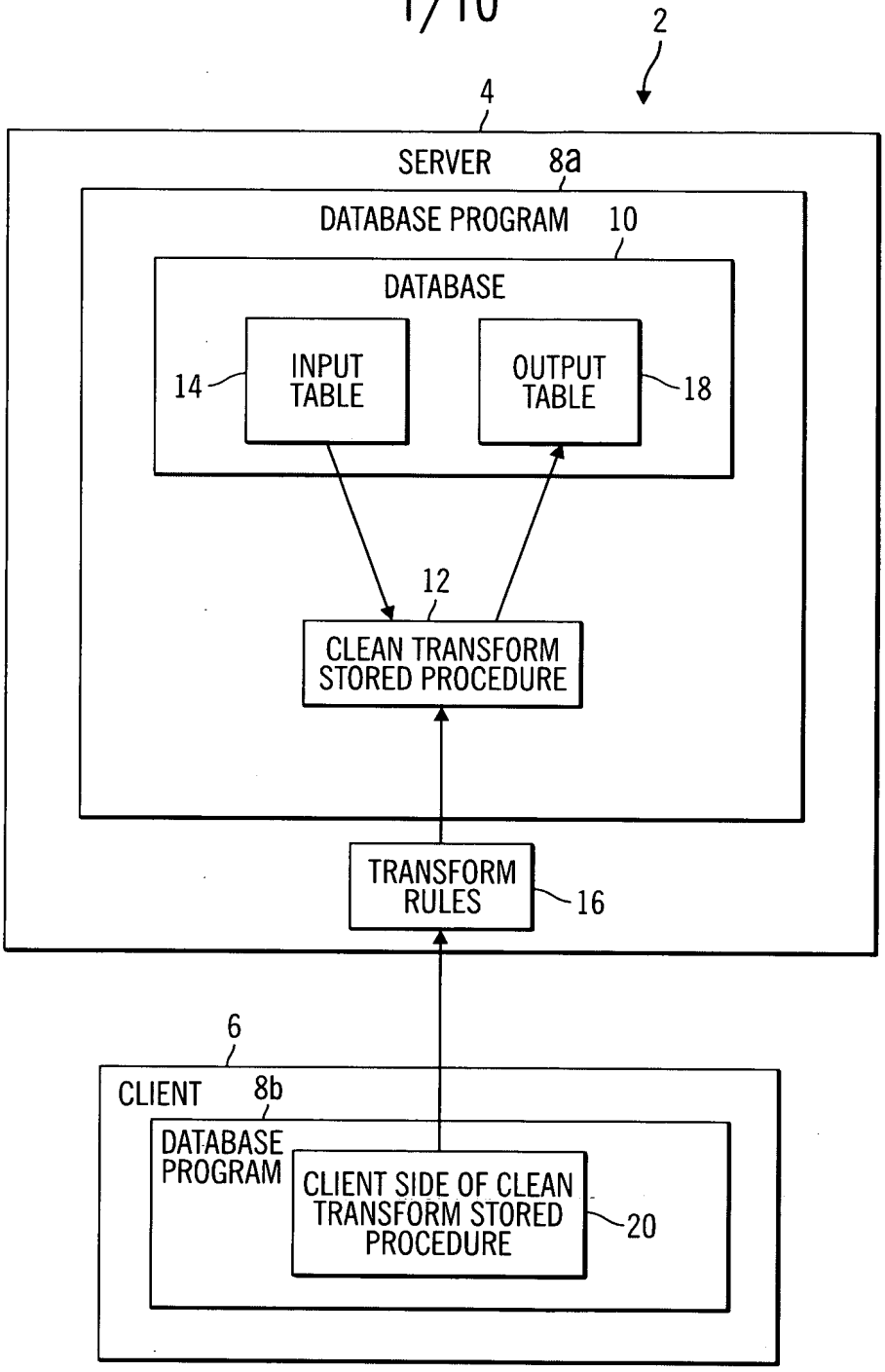

FIG. 1

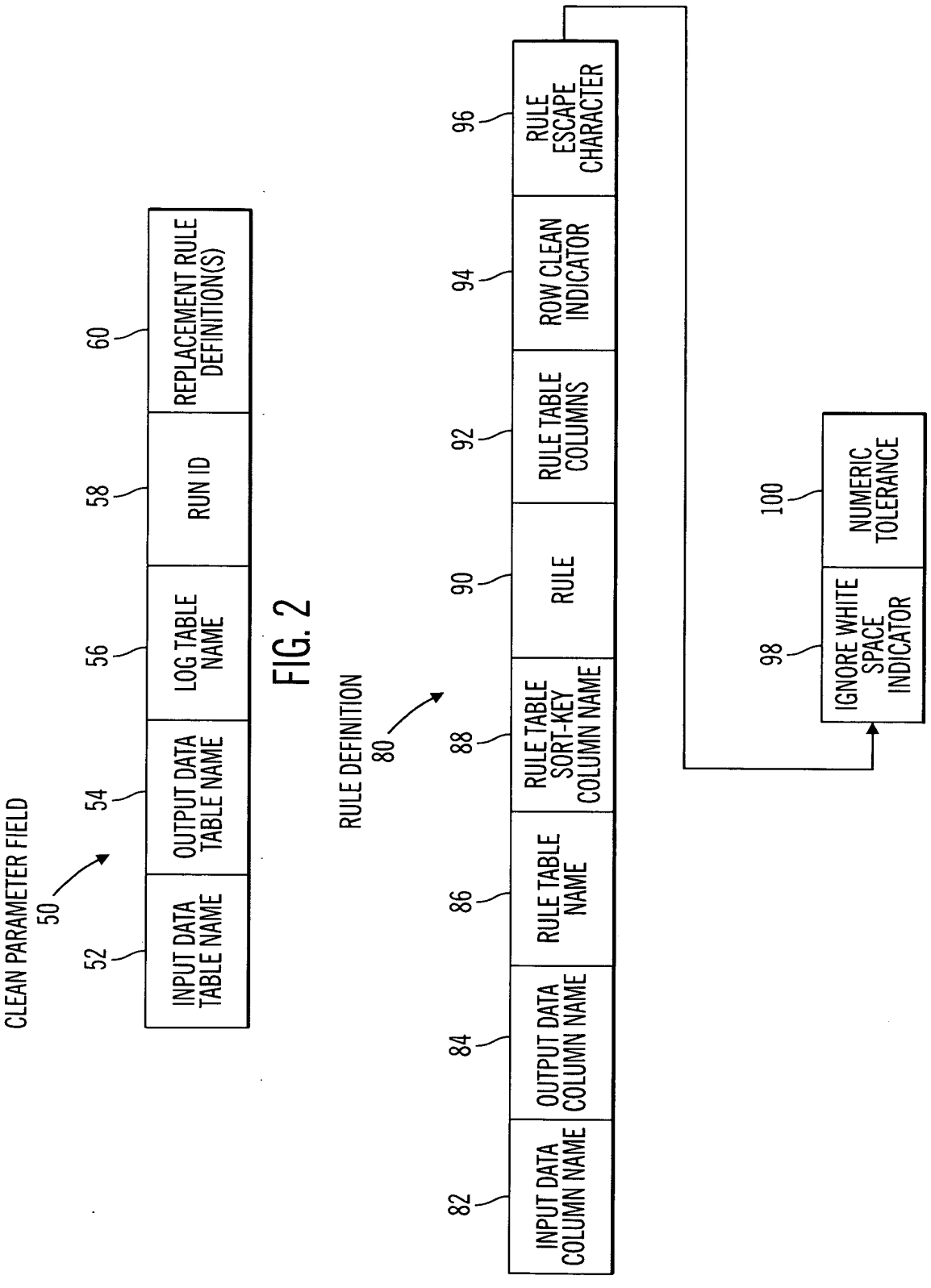
CLEAN PARAMETER FIELD
50
52
INPUT DATA TABLE NAME
54
OUTPUT DATA TABLE NAME
56
LOG TABLE NAME
58
RUN ID
60
REPLACEMENT RULE DEFINITION(S)
RULE DEFINITION
80
82
INPUT DATA COLUMN NAME
84
OUTPUT DATA COLUMN NAME
86
RULE TABLE NAME
88
RULE TABLE SORT-KEY COLUMN NAME
90
RULE
92
RULE TABLE COLUMNS
94
ROW CLEAN INDICATOR
96
RULE ESCAPE CHARACTER
98
IGNORE WHITE SPACE INDICATOR
100
NUMERIC TOLERANCE

FIG. 2

FIG. 3a

FIG. 3b

FIG. 4

FIG. 5

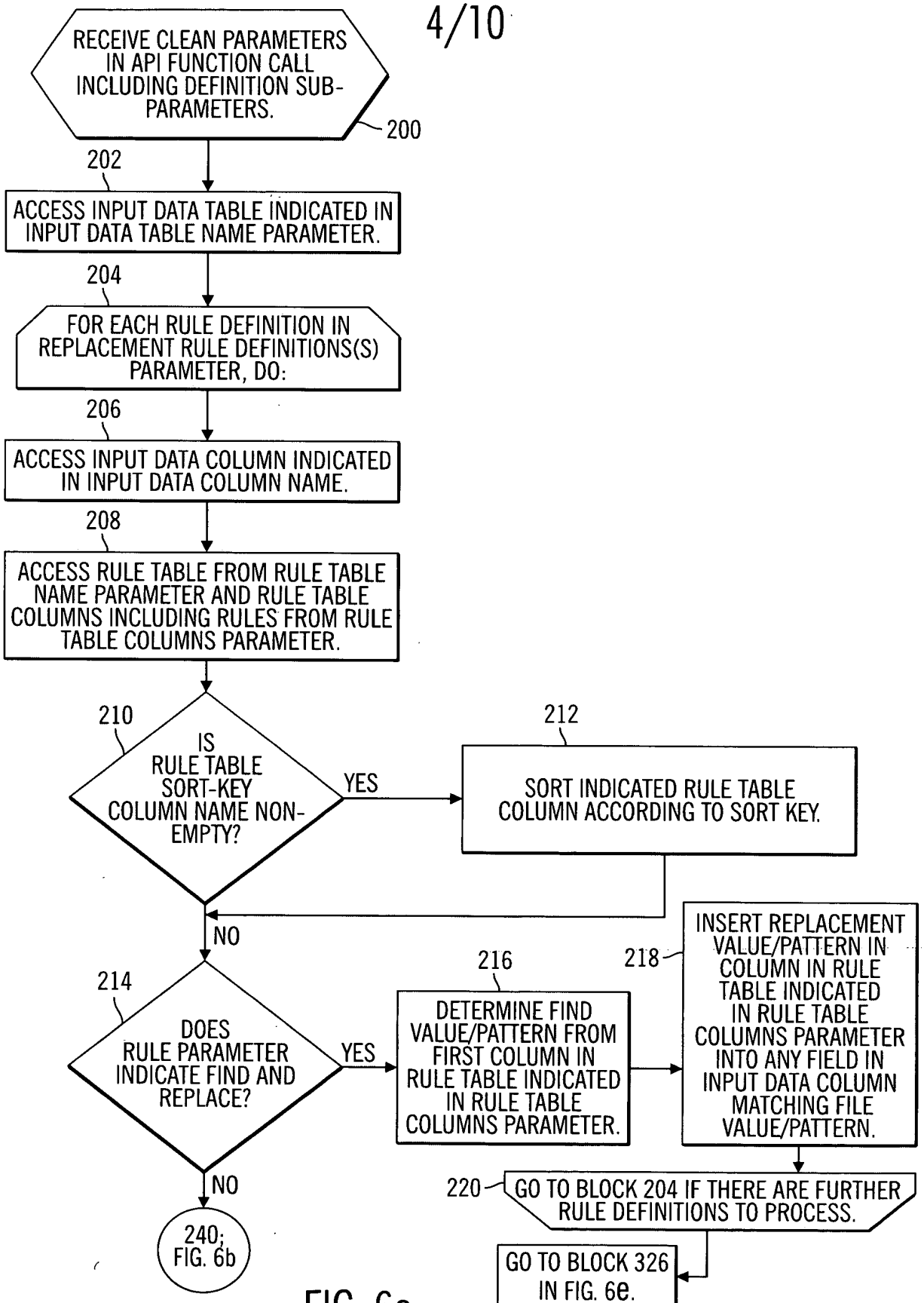
RECEIVE CLEAN PARAMETERS IN API FUNCTION CALL INCLUDING DEFINITION SUB-PARAMETERS.
200
202
ACCESS INPUT DATA TABLE INDICATED IN INPUT DATA TABLE NAME PARAMETER.
204
FOR EACH RULE DEFINITION IN REPLACEMENT RULE DEFINITIONS(S) PARAMETER, DO:
206
ACCESS INPUT DATA COLUMN INDICATED IN INPUT DATA COLUMN NAME.
208
ACCESS RULE TABLE FROM RULE TABLE NAME PARAMETER AND RULE TABLE COLUMNS INCLUDING RULES FROM RULE TABLE COLUMNS PARAMETER.
210
IS RULE TABLE SORT-KEY COLUMN NAME NON-EMPTY?
YES
212
SORT INDICATED RULE TABLE COLUMN ACCORDING TO SORT KEY.
NO
214
DOES RULE PARAMETER INDICATE FIND AND REPLACE?
YES
216
DETERMINE FIND VALUE/PATTERN FROM FIRST COLUMN IN RULE TABLE INDICATED IN RULE TABLE COLUMNS PARAMETER.
218
INSERT REPLACEMENT VALUE/PATTERN IN COLUMN IN RULE TABLE INDICATED IN RULE TABLE COLUMNS PARAMETER INTO ANY FIELD IN INPUT DATA COLUMN MATCHING FILE VALUE/PATTERN.
220
GO TO BLOCK 204 IF THERE ARE FURTHER RULE DEFINITIONS TO PROCESS.
GO TO BLOCK 326 IN FIG. 6e.
NO
240; FIG. 6b

FIG. 6a

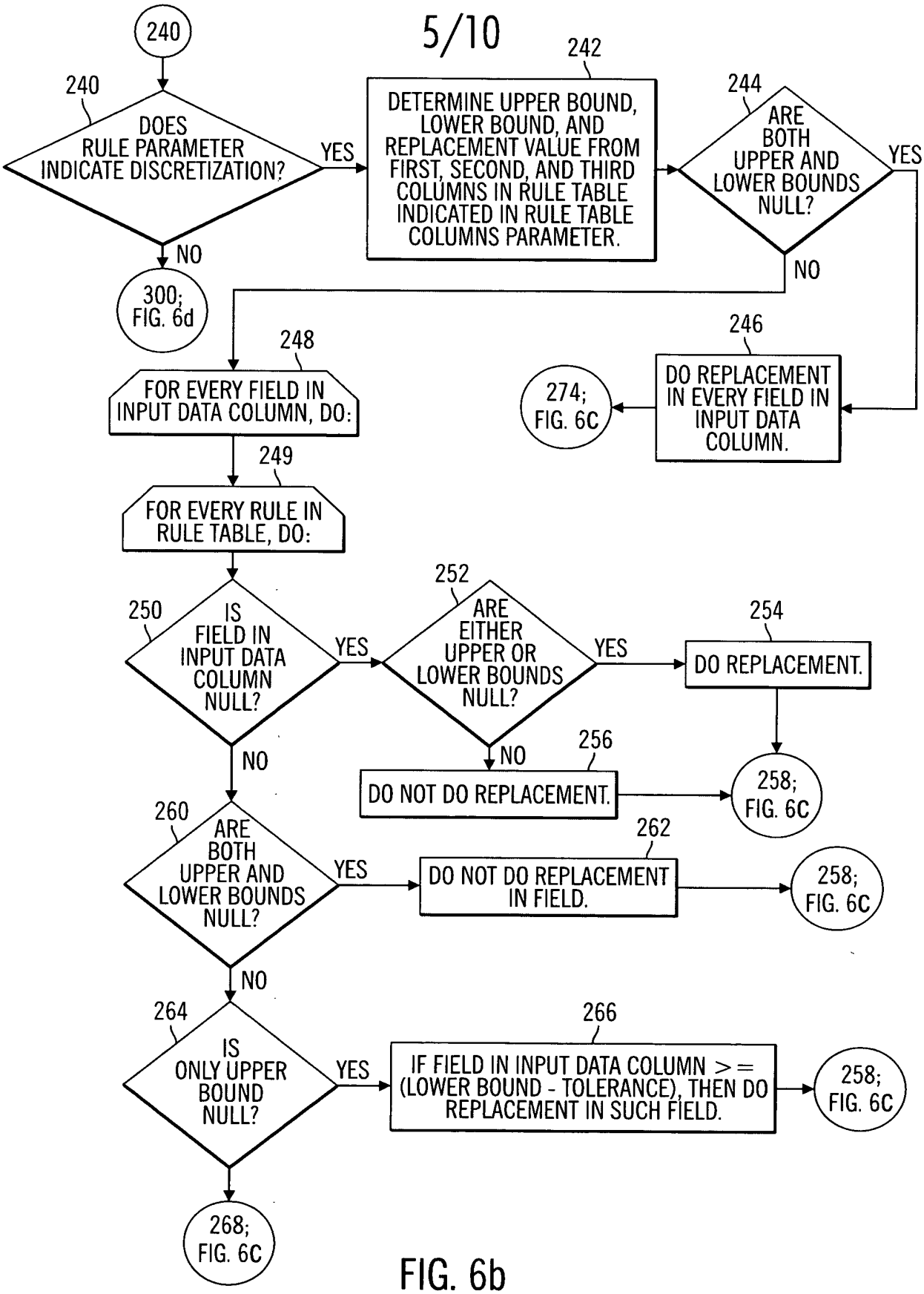
240
240
DOES RULE PARAMETER INDICATE DISCRETIZATION?
YES
NO
300; FIG. 6d
242
DETERMINE UPPER BOUND, LOWER BOUND, AND REPLACEMENT VALUE FROM FIRST, SECOND, AND THIRD COLUMNS IN RULE TABLE INDICATED IN RULE TABLE COLUMNS PARAMETER.
244
ARE BOTH UPPER AND LOWER BOUNDS NULL?
YES
NO
246
DO REPLACEMENT IN EVERY FIELD IN INPUT DATA COLUMN.
274; FIG. 6C
248
FOR EVERY FIELD IN INPUT DATA COLUMN, DO:
249
FOR EVERY RULE IN RULE TABLE, DO:
250
IS FIELD IN INPUT DATA COLUMN NULL?
YES
NO
252
ARE EITHER UPPER OR LOWER BOUNDS NULL?
YES
NO
254
DO REPLACEMENT.
256
DO NOT DO REPLACEMENT.
258; FIG. 6C
260
ARE BOTH UPPER AND LOWER BOUNDS NULL?
YES
NO
262
DO NOT DO REPLACEMENT IN FIELD.
258; FIG. 6C
264
IS ONLY UPPER BOUND NULL?
YES
NO
266
IF FIELD IN INPUT DATA COLUMN >= (LOWER BOUND - TOLERANCE), THEN DO REPLACEMENT IN SUCH FIELD.
258; FIG. 6C
268; FIG. 6C

FIG. 6b

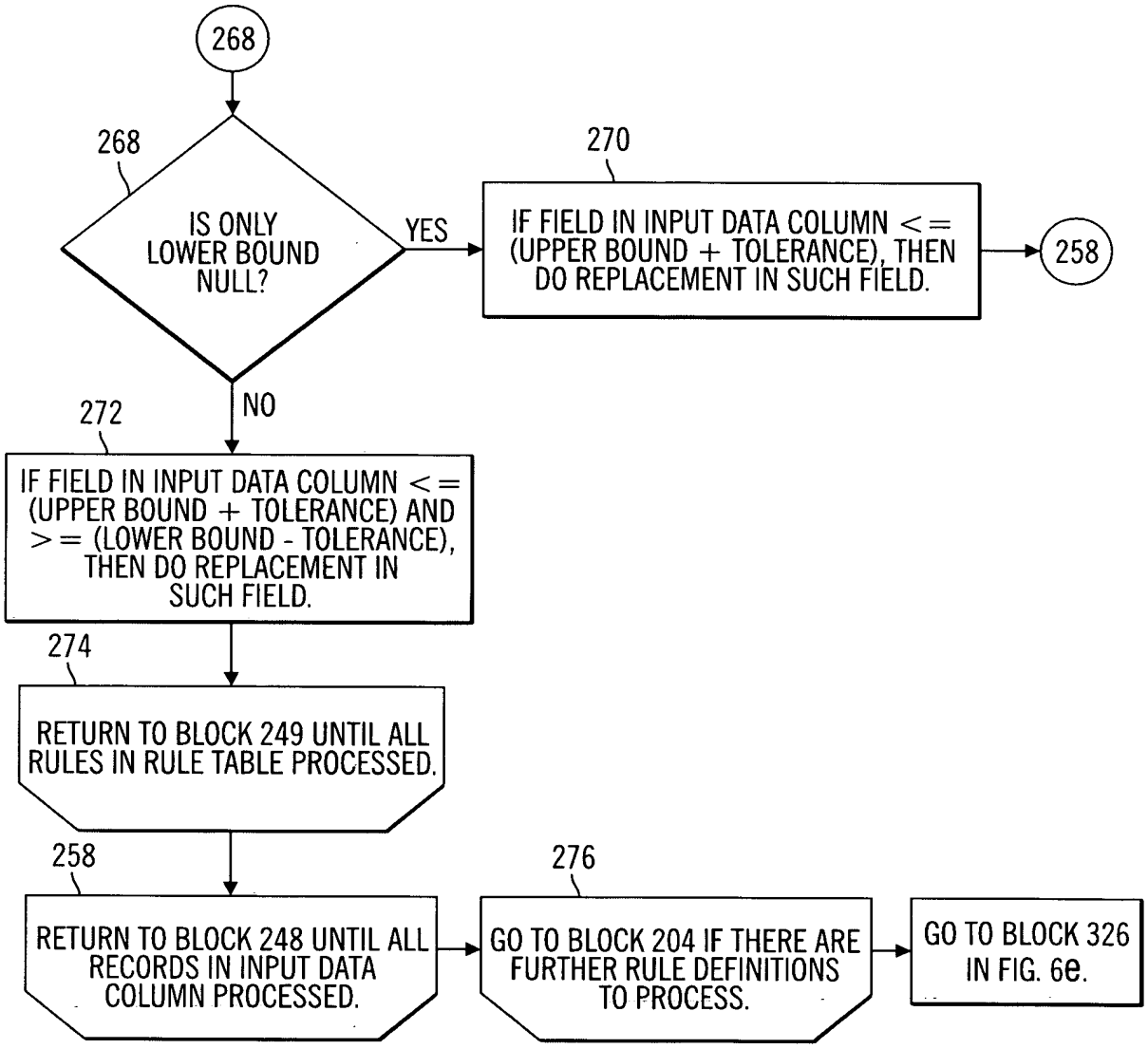
268
268
IS ONLY LOWER BOUND NULL?
YES
270
IF FIELD IN INPUT DATA COLUMN <= (UPPER BOUND + TOLERANCE), THEN DO REPLACEMENT IN SUCH FIELD.
258
NO
272
IF FIELD IN INPUT DATA COLUMN <= (UPPER BOUND + TOLERANCE) AND >= (LOWER BOUND - TOLERANCE), THEN DO REPLACEMENT IN SUCH FIELD.
274
RETURN TO BLOCK 249 UNTIL ALL RULES IN RULE TABLE PROCESSED.
258
RETURN TO BLOCK 248 UNTIL ALL RECORDS IN INPUT DATA COLUMN PROCESSED.
276
GO TO BLOCK 204 IF THERE ARE FURTHER RULE DEFINITIONS TO PROCESS.
GO TO BLOCK 326 IN FIG. 6e.

FIG. 6c

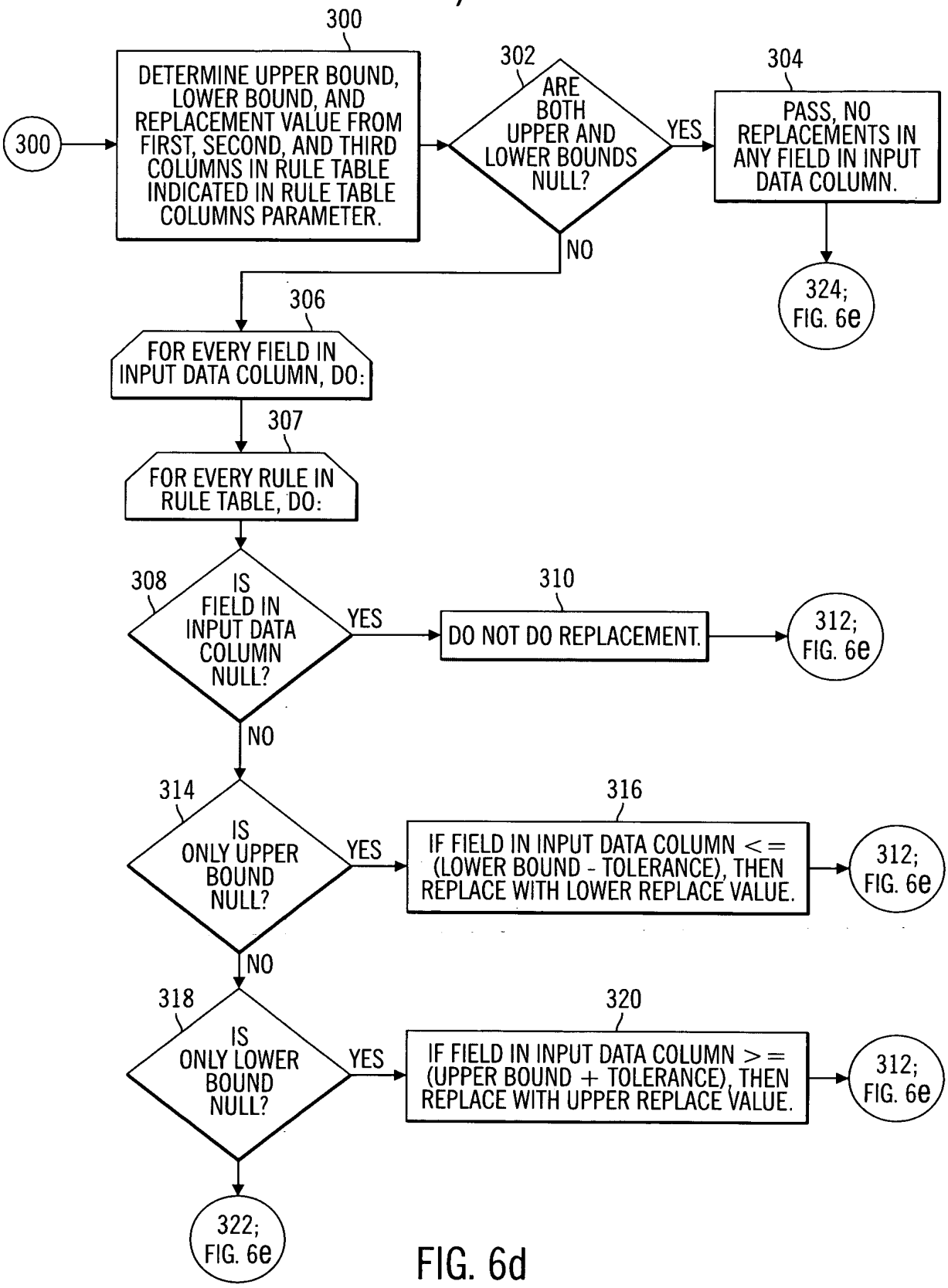
300
DETERMINE UPPER BOUND, LOWER BOUND, AND REPLACEMENT VALUE FROM FIRST, SECOND, AND THIRD COLUMNS IN RULE TABLE INDICATED IN RULE TABLE COLUMNS PARAMETER.
300
302
ARE BOTH UPPER AND LOWER BOUNDS NULL?
YES
304
PASS, NO REPLACEMENTS IN ANY FIELD IN INPUT DATA COLUMN.
324; FIG. 6e
NO
306
FOR EVERY FIELD IN INPUT DATA COLUMN, DO:
307
FOR EVERY RULE IN RULE TABLE, DO:
308
IS FIELD IN INPUT DATA COLUMN NULL?
YES
310
DO NOT DO REPLACEMENT.
312; FIG. 6e
NO
314
IS ONLY UPPER BOUND NULL?
YES
316
IF FIELD IN INPUT DATA COLUMN <= (LOWER BOUND - TOLERANCE), THEN REPLACE WITH LOWER REPLACE VALUE.
312; FIG. 6e
NO
318
IS ONLY LOWER BOUND NULL?
YES
320
IF FIELD IN INPUT DATA COLUMN >= (UPPER BOUND + TOLERANCE), THEN REPLACE WITH UPPER REPLACE VALUE.
312; FIG. 6e
322; FIG. 6e

FIG. 6d

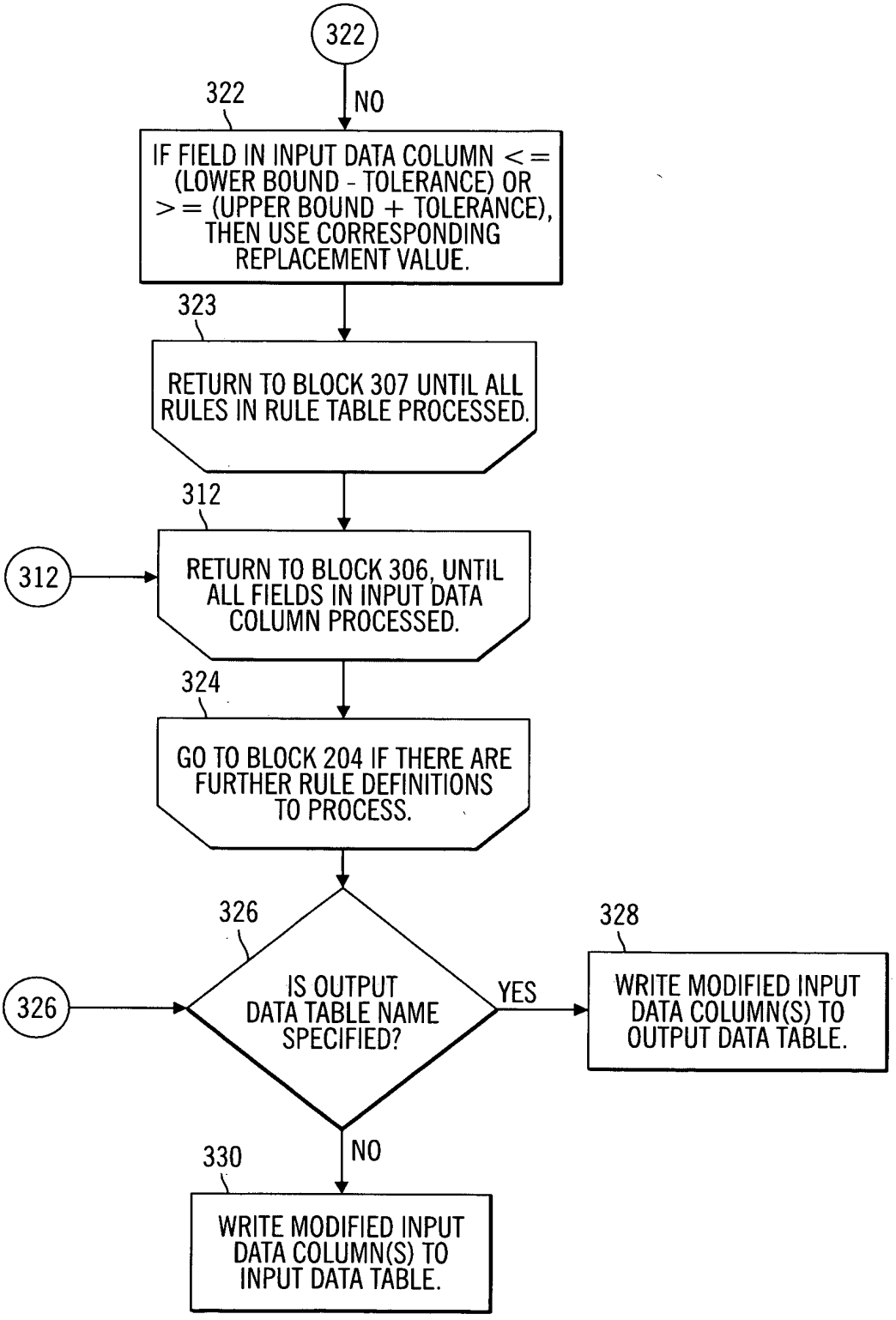
322
322
NO
IF FIELD IN INPUT DATA COLUMN <= (LOWER BOUND - TOLERANCE) OR >= (UPPER BOUND + TOLERANCE), THEN USE CORRESPONDING REPLACEMENT VALUE.
323
RETURN TO BLOCK 307 UNTIL ALL RULES IN RULE TABLE PROCESSED.
312
312
RETURN TO BLOCK 306, UNTIL ALL FIELDS IN INPUT DATA COLUMN PROCESSED.
324
GO TO BLOCK 204 IF THERE ARE FURTHER RULE DEFINITIONS TO PROCESS.
326
326
IS OUTPUT DATA TABLE NAME SPECIFIED?
YES
328
WRITE MODIFIED INPUT DATA COLUMN(S) TO OUTPUT DATA TABLE.
NO
330
WRITE MODIFIED INPUT DATA COLUMN(S) TO INPUT DATA TABLE.

FIG. 6e

CITY	PRODUCT NAME	PRODUCT CATEGORY	SALES ($)	UNIT SALES
VARCHAR (20)	VARCHAR (30)	VARCHAR (15)	DECIMAL (8,2)	INTEGER
11%LAX455	PICK AXE	IATE	5897.32	1100
SJC1298%SEA	SMALL SHOVEL	JXJ8972	456.98	43
%LAX6%SF094	TARP	HAU1245	893.12	422
%SJC34LAX45	FLARES	AX111	1200.45	1189
4515%SEA	NULL	AC183	333.9	23
3452%SJC34	""	MIS222	1818.4	15002

FIG. 7

FIND VALUE	REPLACE VALUE
VARCHAR (10)	VARCHAR (20)
%%LAX%	LOS ANGELES
%%DFW%	DALLAS/FORT WORTH
%%SJC%	SAN JOSE, CA
%%SFO%	SAN FRANCISCO
%%SEA%	SEATTLE/TACOMA

FIG. 8a

LOWER BOUND	LOWER REPLACEMENT	UPPER BOUND	UPPER REPLACEMENT
INTEGER	INTEGER	INTEGER	INTEGER
100	100	15000	NULL

FIG. 8b

SORT COLUMN	LOWER BOUND	UPPER BOUND	REPLACE VALUE
INTEGER	VARCHAR (6)	VARCHAR (6)	VARCHAR (10)
1	I	IZZZZ	SPECIAL
2	HA	JZ	HARDWARE
3	LA	NZ	MISCEL
4	AA	BJ	WINTER STORM
5	RX	NULL	WINTER STORM

FIG. 8c

FIND VALUE	REPLACE VALUE
VARCHAR (5)	VARCHAR (10)
NULL	""

FIG. 8d